

**Request to Archive
With The National Centers for Environmental Information
For Big Earth Data Initiative For Chromatograph for Atmospheric Trace Species Program
Provided by NOAA/OAR/ESRL**

2017-03-21

This information will be used by NCEI to conduct an appraisal and make a decision on the request.

1. Who is the primary point of contact for this request?

Geoffrey Dutton

DOC/NOAA/ESRL/GMD > Global Monitoring Division, Earth System Research Laboratory, NOAA, U.S.

Department of Commerce

Affiliate

303-497-6086

geoff.dutton@noaa.gov

email

2. Name the organization or group responsible for creating the dataset.

DOC/NOAA/OAR/ESRL/GMD/HATS > Halocarbons and other Atmospheric Trace Species Group, Global Monitoring Division, Earth System Research Laboratory, OAR, NOAA, U.S. Department of Commerce

3. Provide an overview summarizing the scope of data you want to archive. Describe the outputs, data variables, including their measurement resolution and coverage.

Hourly, daily medians, monthly medians, hemispherical and global means of 8 halocarbons as acquired by the Chromatograph for Atmospheric Trace Species (CATS) insitu instruments. The CATS equipment is installed at six locations.

4. What is the time period covered by the dataset? (YYYY-MM-DD, YYYY-MM or YYYY)

From 1998

Ongoing as continuous updates to the data record

5. Edition or version number(s) of the dataset:

20170302

6. Approximate date when the dataset was or will be released to the public:

2017-03-02

7. Who are the expected users of the archived data? How will the archived data be used?

Scientist

8. Has the dataset undergone user evaluation and/or an independent review process? Did NCEI participate in design reviews?

No

9. Describe the dataset's relationship to other archived datasets, such as earlier versions or related source data. If this is a new version, how does it improve upon the previous version(s)?

This is a new version compiled from text files on the NOAA/ESRL/GMD ftp site.

10. List the input datasets and ancillary information used to produce the data.

The original datasets are located at: <ftp://ftp.cmdl.noaa.gov/hats>

11. List web pages and other links that provide information on the data.

<https://www.esrl.noaa.gov/gmd/hats/insitu/cats/>

12. List the kinds of documents, metadata and code that are available for archiving. For example, data format specifications, user guides, algorithm documentation, metadata compliant with a standard such as ISO 19115, source code, platform/instrument metadata, data/process flow diagrams, etc.

1. Platform/instrument metadata is included in files' metadata.

13. Indicate the data file format(s).

1. netCDF-4

14. Are the data files compressed?

netCDF-4/HDF5 compression

15. Provide details on how the files are named and how they are organized (e.g., file_name_pattern_YYYYMM.tar in monthly aggregations).

There is a file for each molecule the CATS instruments measure. Within the netCDF4 file are four groups specifically for data rates (hourly, daily medians, monthly medians, and hemispheric and global medians).

16. Explain how to access sample data files and/or a file listing for previewing. If it is not available now, when will it be available?

Example of original text data files: <ftp://ftp.cmdl.noaa.gov/hats/cfcs/cfc11/insituGCs/CATS/monthly/>

17. What is the total data volume to be submitted?

Historic Data: all historic data or data submitted as a completed collection.

Total Data Volume: 100MB

Number of Data Files: 8

18. Are later updates, revisions or replacement files anticipated? If so, explain the conditions for submitting these additional data to the archive.

The measurements are ongoing.

19. Describe the server that will connect to the ingest server at NCEI for submitting the data.

Physical Location: Boulder, Colorado

System Name: n/a

System Owner: DOC/NOAA/ESRL/GMD > Global Monitoring Division, Earth
System Research Laboratory, NOAA, U.S. Department of
Commerce

Additional Information: No provisioned server

20. What are the possible methods for submitting the data to NCEI? Select all that apply.

1. FTP PULL
2. FTP PUSH
3. SFTP PUSH

21. Identify how you would like NCEI to distribute the data. Web access support depends on the resources available for the dataset.

1. Direct download links

22. Will there be any distribution, usage, or other restrictions that apply to the data in the archive?

Constraint Type	Description
Use	If you use this data in a publication or report, we would appreciate that you contact the principal investigators (PIs) first and discuss your interests. We are trying to encourage scientific discussion. The PIs can discuss the quality of the data and any potential problems with its analysis. Recent data (less than 1.5 years) are considered preliminary. Please contact the PIs for more information.

23. Discuss the rationale for archiving the dataset and the anticipated benefits. Mention any risks associated with not archiving the dataset at NCEI.

Wider distribution and discovery of GMD CATS data.

24. Are the data archived at another facility or are there plans to do so? Please explain.

On the NOAA/ESRL/GMD ftp site.

25. Is there an existing agreement or requirement driving this request to archive? Have you already contacted someone at NCEI?

BEDI and GMD agreement to archive data.

26. Do you have a data management plan for your data?

No

27. Have funds been allocated to archive the data at NCEI?

Yes by the Big Earth Data Initiative

28. Identify the affiliated research project, its sponsor, and any project/grant ID as applicable.

N/A

29. Is there a desired deadline for NCEI to archive and provide access to the data?

Archive by: 2017-10-01

Accessible by:

30. Add any other pertinent information for this request.

This is a test submission.

